

画像認識能力の獲得の中核である、クラスタリング技術について考えて行ってみます。

画像はピクセルの2次元マップです。この解像度を下げていくのに、長さや方向が一定の量子化された線分というものを得て、これを点とするマップで、新しく画像を造りスタックしていくのがクラスタリングでのデータ収集です。

枠と枠に所属する部品群という構造を捉えて、枠をマップとして切り取り、重心を一致させて重ね合わせていきます。

枠の最も外側のものから、内部の枠まで、同じものの違った形のデータであるか、異なるもののデータであるかというものの管理が重要です。人間というもつとも外側の枠は、犬というもつとも外側の枠に対向します。画像の違いは、枠の中で、どこが違うのかということを知ることが重要であることが分かります。もつとも外側の人間のなかの内側の枠である鼻は犬の鼻と比較されます。そのようなことから、画像の重ね合わせは、交差法で、グラフ構造の画像枠群のネットワークを造ることが必須であることが分かります。鼻とか、目とかの枠毎に画像の重ね合わせをしていくのです。

画像の重ね合わせを行い違いを検出する技術をここではクラスタリングと言っています。前記の手法でスタックした画像は、種類毎に分類されていきますから、種類毎の画像の変位が分かります。こういう傾向にあるというような塊というか、点の分布に成るはずで、人間の鼻なら鼻の画像の重なりができるのです。犬なら犬の鼻の画像と分類できます。そのとき、2つの種類の鼻のどこかに互いに特徴的に異なる部分が見えるはずで、その違いを捉えたい・・・どうすれば良いでしょうか。

画像の重心を重ね合わせると前に記述しました。ここで、重心を極座標の原点として例えば16方向に向かい、2種類の画像群の造る点の分布の重なり具合を見るということも解になることが分かります。分布する各点がガウス分布するポテンシャルを持っていて、そのポテンシャルの重ね合わせをもって一つの2次元分布関数ができます。その尤も高い点を1として正規化すると、2つの分布関数ができます。その関数の場所による引き算をすると、1から-1までの重ね合わせの状況が1直線上に見えるようになります。これによって、人間の鼻の特徴の犬の鼻との特徴の違いというのがマップで得ることになります。

ここで、枠の認知について考えて見ます。目だけで顔の輪郭が明示されなくても、輪郭は想定されます。これは、目からのパーセプトロンによる評価システムのおかげです。このように枠を推定すること・・・線分分布を推定することも必要になります。そうして、誤った場合は事象も少ないでしょうから、誤りデータ処理のように重み付き重なりという評価を得て、学習データとすることです。

おわり